

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problems Mailbox.**

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-320327

(43) 公開日 平成10年(1998)12月4日

(51) Int.Cl.⁶

G 0 6 F 13/00

H 0 4 L 12/28

識別記号

3 5 3

F I

G 0 6 F 13/00

H 0 4 L 11/00

3 5 3 T

3 1 0 D

審査請求 未請求 請求項の数17 O L (全 16 頁)

(21) 出願番号 特願平10-40139

(22) 出願日 平成10年(1998)2月23日

(31) 優先権主張番号 特願平9-66717

(32) 優先日 平9(1997)3月19日

(33) 優先権主張国 日本 (J P)

(71) 出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中4丁目1番
1号

(72) 発明者 山本 眞二

神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内

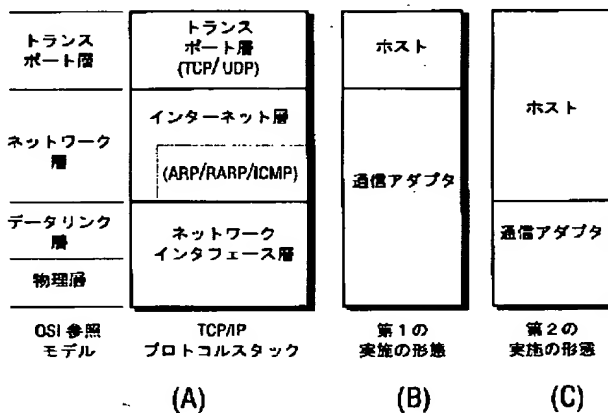
(74) 代理人 弁理士 服部 毅巖

(54) 【発明の名称】 二重化された通信アダプタの切替方法、切替方式、および切替用プログラムを格納した記録媒体

(57) 【要約】 (修正有)

【課題】 ホストの通信アダプタが故障した場合に、予備用の通信アダプタに速やかに切り替えて通信を継続できるようにする。

【解決手段】 ホストに第1と第2の通信アダプタを接続し、ホットスタンバイシステムを構成する。ホストが第1の通信アダプタの故障を検出した時に、第1から第2の通信アダプタに切替え、ネットワークテーブル(ARPキャッシュ)を初期化し、第2のホストへの制御メッセージの送信を起動する。第2の通信アダプタは、ARP要求メッセージを同報送信し、第2のホストからARP応答メッセージを受信し、ARP応答メッセージに含まれるアドレス情報に基づいて、前記制御メッセージを第2のホストに送信し、前記制御メッセージに対する第2のホストからの応答メッセージを受信する。これによって第1のホストは、第1から第2の通信アダプタへの切り替えが終了したと認識する。



【特許請求の範囲】

【請求項1】第1のホストの通信機能が第1と第2の通信アダプタによって二重化され、第2のホストを含む複数のホスト間の通信をTCP/IPプロトコルを使用して行なうネットワークシステムで、第1の通信アダプタから第2の通信アダプタに切り替える切替方法において、

- (a) 前記第1のホストが前記第1の通信アダプタの故障を検出した時に、前記第1の通信アダプタを切り離し、
 - (b) 前記第2の通信アダプタを活性化し、
 - (c) 前記第1のホストから前記第2のホストへの制御メッセージの送信を起動し、
 - (d) 前記第2の通信アダプタからARP 要求メッセージを同報送信し、
 - (e) 前記第2の通信アダプタが前記第2のホストから受信したARP 応答メッセージに含まれるアドレス情報に基づいて、前記制御メッセージを前記第2のホストに送信し、
 - (f) 前記制御メッセージに対する前記第2のホストからの応答メッセージを受信し、
 - (g) 前記応答メッセージの受信をもって、前記第1の通信アダプタから前記第2の通信アダプタへの切り替えが終了したと認識するようにした、
- ことを特徴とする切替方法。

【請求項2】前記第1のホストはトランスポート層の通信プロトコルを制御し、前記第1と第2の通信アダプタはインターネット層の通信プロトコルを制御することを特徴とする請求項1記載の切替方法。

【請求項3】前記ステップ(c) で送信起動される前記制御メッセージは、ICMPプロトコルのエコー要求メッセージであることを特徴とする請求項1記載の切替方法。

【請求項4】前記ステップ(b) は、前記第2の通信アダプタのネットワークテーブルを初期化することを含むことを特徴とする請求項1記載の切替方法。

【請求項5】第1のホストの通信機能が第1と第2の通信アダプタによって二重化され、第2のホストを含む複数のホスト間の通信をTCP/IPプロトコルを使用して行なうネットワークシステムで、前記第1の通信アダプタから前記第2の通信アダプタに切り替える切替方法において、

- (a) 前記第1のホストが前記第1の通信アダプタの故障を検出した時に、前記第1の通信アダプタを切り離し、
 - (b) 前記第2の通信アダプタを活性化し、
 - (c) 前記第1のホストから前記第2の通信アダプタを経由して前記第2のホストにARP 応答メッセージを送信する、
- ことを特徴とする切替方法。

【請求項6】前記第1のホストはインターネット層の通信プロトコルを制御し、前記第1と第2の通信アダプタはネットワークインタフェース層の通信プロトコルを制御することを特徴とする請求項5記載の切替方法。

【請求項7】前記ステップ(c) で送信起動されるARP 応答メッセージは、その送信元アドレスフィールドに前記第1のホストのIPアドレスと前記第2の通信アダプタのMACアドレスを含むことを特徴とする請求項5記載の切替方法。

【請求項8】第1のホストの通信機能が第1と第2の通信アダプタによって二重化され、第2のホストを含む複数のホスト間の通信をTCP/IPプロトコルを使用して行なうネットワークシステムで、ホストの通信アダプタを故障時に切り替える切替方式において、

- (a) 前記第1の通信アダプタの故障を検出する故障検出手段と、
 - (b) 前記故障検出手段によって検出された故障に応答して、前記第1の通信アダプタを切り離し、前記第2の通信アダプタを活性化する活性化手段と、
 - (c) 前記活性化手段によって活性化された前記第2の通信アダプタを経由し前記第1のホストから前記第2のホストへの制御メッセージの送信を起動する制御メッセージ送信手段と、
 - (d) 前記制御メッセージに対する前記第2のホストからの応答メッセージの受信をもって、前記第1の通信アダプタから前記第2の通信アダプタへの切り替えが終了したと認識する完了認識手段と、
- を備えたことを特徴とする切替方式。

【請求項9】前記第2の通信アダプタは、制御メッセージ送信手段による前記制御メッセージの送信要求に応答して、ARP 要求メッセージを同報送信するARP 要求手段と、

前記第2のホストから受信したARP 応答メッセージに含まれるアドレス情報に基づいて、前記制御メッセージを前記第2のホストに送信する送信手段と、

を備えたことを特徴とする請求項8記載の切替方式。

【請求項10】前記第1のホストはトランスポート層の通信プロトコルを制御し、前記第1と第2の通信アダプタはインターネット層の通信プロトコルを制御することを特徴とする請求項8記載の切替方式。

【請求項11】前記制御メッセージは、ICMPプロトコルのエコー要求メッセージであることを特徴とする請求項8記載の切替方式。

【請求項12】第1のホストの通信機能が第1と第2の通信アダプタによって二重化され、第2のホストを含む複数のホスト間の通信をTCP/IPプロトコルを使用して行なうネットワークシステムで、第1のホストの通信アダプタを故障時に切り替える切替方式において、

- (a) 前記第1の通信アダプタの故障を検出する故障検出手段と、
- (b) 前記故障検出手段によって検出された故障に応答して、前記第1の通信アダプタを切り離し、前記第2の通信アダプタを活性化する活性化手段と、
- (c) 前記第1のホストから前記第2のホストへのARP 応

答メッセージを、前記活性化手段によって活性化された前記第2の通信アダプタを経由して送信するARPメッセージ送信手段と、

を備えたことを特徴とする切替方式。

【請求項13】前記第1のホストはインターネット層の通信プロトコルを制御し、前記第1と第2の通信アダプタはネットワークインタフェース層の通信プロトコルを制御することを特徴とする請求項12記載の切替方式。

【請求項14】前記ARPメッセージ送信手段が送信する前記ARP応答メッセージは、その送信元アドレスフィールドに前記第1のホストのIPアドレスと前記第2の通信アダプタのMACアドレスを含むことを特徴とする請求項12記載の切替方式。

【請求項15】前記第1のホストは、前記第1と第2の通信アダプタを含む3台以上の通信アダプタを有し、前記第1の通信アダプタを含む2台以上を現用通信アダプタとして使用し、前記第2の通信アダプタを含む残りをバックアップ用通信アダプタとした、ことを特徴とする請求項8または請求項12記載の通信アダプタの切替方式。

【請求項16】第1のホストの通信機能が第1と第2の通信アダプタによって二重化され、第2のホストを含む複数のホスト間の通信をTCP/IPプロトコルを使用して行なうネットワークシステムで、第1のホストが第1の通信アダプタから第2の通信アダプタへの切替を制御するためのプログラムを記録した媒体において、

(a) 前記第1の通信アダプタの故障を検出した時に、前記第1の通信アダプタを切り離し、

(b) 前記第2の通信アダプタを活性化すると共に、前記第2の通信アダプタ内のネットワークテーブルを初期化させ、

(c) 前記第2の通信アダプタにARP要求メッセージを送信させるべく、前記第1のホストから前記第2のホストへの制御メッセージの送信を、前記第2の通信アダプタに要求し、

(d) 前記ステップ(c)の結果として前記第2のホストに送信された前記制御メッセージに対する前記第2のホストからの応答メッセージを受信し、

(e) 前記応答メッセージの受信をもって、前記第1の通信アダプタから前記第2の通信アダプタへの切り替えが終了したと認識する、

ことを第1のホストを構成するコンピュータに実行させるためのプログラムを格納した記録媒体。

【請求項17】第1のホストの通信機能が第1と第2の通信アダプタによって二重化され、第2のホストを含む複数のホスト間の通信をTCP/IPプロトコルを使用して行なうネットワークシステムで、第1のホストが第1の通信アダプタから第2の通信アダプタへの切替を制御するためのプログラムを記録した媒体において、

(a) 前記第1の通信アダプタの故障を検出した時に、前記第1の通信アダプタを切り離し、

(b) 前記第2の通信アダプタを活性化すると共に、前記第1のホスト自身が有するネットワークテーブルを更新し、

(c) 活性化された前記第2の通信アダプタを経由して前記第2のホストにARP応答メッセージを送信する、ことを第1のホストを構成するコンピュータに実行させるためのプログラムを格納した記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は耐故障性を高めたホストの通信機能をサポートする二重化された通信用アダプタ装置の切替方法と切替方式、および前記切替方法を実行するプログラムを記録した媒体に関し、特にTCP/IPプロトコルを用いたホスト間通信をサポートする二重化された通信アダプタにおいて、障害発生時に通信を切断することなく切り替えを行なうための切替方法と切替方式、および前記切替方法を実行するプログラムを記録した媒体に関する。

【0002】

【従来の技術】TCP/IP(Transmission Control Protocol/Internet Protocol)プロトコルはコンピュータネットワークにおいて広く使われている通信手順である。従来、このTCP/IPプロトコルを用いてデータの送受信を行うホスト間のネットワークシステムがある。各ホストにはTCP/IPプロトコルスタックの機能を提供する通信アダプタが、周辺装置の一つとしてローカルチャネルあるいはシステムバスを通じて結合される。サーバーシステムなど、通信機能に関して特に高い信頼性が要求される場合においては、通信アダプタを二重化して、ホットスタンバイシステムを構成する。これは現用とバックアップ用の二つの通信アダプタを用意し、現用通信アダプタに異常が生じてバックアップ用通信アダプタに自動的に切り替えることによって、ダウンタイムを減らすためのものである。

【0003】図7はこのようなホスト間通信システムの構成を示す図である。ホスト10は二重化された通信アダプタ11と12を通してLAN 100に接続されている。通信アダプタ11と12は、両者とも常にネットワーク上のエンティティとして動作しているが、ホスト10は通常は専ら通信アダプタ11を通して通信を行ない、もう一方の通信アダプタ12は待機状態にある。LAN 100には別のホスト20が通信アダプタ21を通して接続されている。またLAN 100はルータ30を経由して別のLAN 200にリンクしている。LAN 200には多くのコンピュータ210、220等が接続されている。

【0004】ネットワークで相互接続されたこれらの装置は、TCP/IPプロトコルを使用して互いにデータ通信を行なう。このため各装置には一意なアドレスが割り当て

られている。例えばホスト10、20にはIPアドレス“A”、“B”がそれぞれ設定されている。これらは論理アドレスと呼ばれるものである。実際に通信を行なうには、プロトコルの下位層の通信を制御するための物理アドレス、すなわちMAC アドレスがさらに必要となる。これらは下位層の通信を制御する通信アダプタに割り当てられるものである。例えばホスト10の通信アダプタ11、12にはMAC アドレス“a”、“c”がそれぞれ割り当てられている。同様にホスト20の通信アダプタ21にはMACアドレス“b”が割り当てられている。

【0005】IPアドレスとMAC アドレスの対応関係は、後述のARP プロトコルによって知ることができる。またARP プロトコルによって得られたアドレス情報は、通信アダプタあるいはホストの中にあるネットワークテーブル(ARPキャッシュ)に管理されている。

【0006】一般にTCP/IPプロトコルを使用した通信では、通信経路の途中のルータが異常になった場合でも、一定時間を経過した後、別のルータを経由して通信を継続することが可能である。

【0007】またホストの現用通信アダプタが異常となった場合には、バックアップ用の通信アダプタへの切り替えが自動的に行われる。しかし従来はこの切り替えを瞬時に行なうことはできず、一定の時間(例えば20分)を要していた。この問題について以下に詳しく説明する。

【0008】まず図8を用いて、従来の通信システムの動作を説明する。図8は、図7で説明したネットワークにおいて通信アダプタを切り替える処理を示すシーケンス図である。ホスト10と20は以下のステップに従って通信を行なう。

【0009】[S71] ホスト10が通信アダプタ11に対してデータの送信依頼を行う。具体には、下記のメッセージが通信アダプタ11に与えられる。

“ [B] [A] [データ] ”

ここで、先頭のフィールド“B”は宛先IPアドレスであり、第2番目のフィールド“A”は発信元IPアドレスである。第3番目のフィールドはペイロード、すなわち実際に相手に送信するデータである。本明細書中のメッセージフォーマットは、説明をわかりやすくするために簡略化されている。実際に実施する場合は、TCP/IPの標準プロトコルの規定に従う。

【0010】[S72] 宛先のMAC アドレス“b”があらかじめ分かっているならば、通信アダプタ11はそれをメッセージに付加してネットワークに送信できる。しかし初期段階では、通信アダプタ11のネットワークテーブル11aには宛先IPアドレス“B”に対応するMAC アドレスが登録されていないので、すぐにデータを送ることができない。そこで通信アダプタ11は一旦このデータ送信を保留し、相手のMAC アドレスを調べるために下記のARP 要求をネットワーク上のすべてのノードに対して送信する。なお

ARP とはAddress Resolution Protocol の略語であり、与えられたIPアドレスに対応するMAC アドレスを得るためのプロトコルである。

【0011】“ [同報] [a] [B] [A] ”

ここで、先頭のフィールド“ [同報] ”は全てのMAC アドレスに送信することを表すコードである。第2番目のフィールド“ [a] ”は発信元のMAC アドレス、すなわち現用通信アダプタ11のMAC アドレスである。第3番目のフィールド“ [B] ”は宛先IPアドレスを表し、第4番目のフィールド“ [A] ”は発信元IPアドレスを表す。

【0012】[S73] バックアップ用通信アダプタ12は待機中ではあるが、ネットワークのアクティビティを常時監視しており、アドレス更新に関する情報があれば、それを取り込むように構成されている。従って通信アダプタ12は、ステップS72で送信されたARP 要求の情報を取り込み、自分自身のネットワークテーブル12aに登録する。ここではホスト10のIPアドレス“A”とMAC アドレス“a”の対応関係がネットワークテーブル12aに登録される。このテーブルエントリを次のように表わす。

【0013】(c | a:A)

最初の“c”はMAC アドレス“c”の装置のネットワークテーブルを意味する。2番目の“a”は受信したMAC アドレスを、また3番目の“A”は対応するIPアドレスを意味している。本明細書ではネットワークテーブルのエントリはこのような形式で表わす。ただし図面においてはグラフィカルな形式で表わす。

【0014】[S74] 通信アダプタ21は、ステップS72で送信されたARP 要求を受信してホスト20に渡す。またホスト20は、自分自身のネットワークテーブルにホスト10のMAC アドレス“ [a] ”とIPアドレス“A”の対応関係を表わすエントリ(b | a:A)に登録する。

【0015】[S75] ホスト20は、前記のARP 要求メッセージが自分宛であることを認識、通信アダプタ21を経由してARP 応答メッセージをホスト10に返信する。

[S76] 通信アダプタ11は、ステップS75で返信されたARP 応答を受信し、ネットワークテーブル11aに、ホスト20のMAC アドレス“ [b] ”とIPアドレス“B”の対応関係を表わすエントリ(a | b:B)に登録する。

【0016】[S77] 宛先MAC アドレスを得ることができたので、通信アダプタ11はステップS72で保留しておいたデータをホスト20に向けて送信する。この時のメッセージには下記のようにIPアドレス、MAC アドレスを設定する。

【0017】

“ [b] [a] [B] [A] [データ] ”

ここで、先頭のフィールド“ [b] ”は宛先MAC アドレス、第2番目のフィールド“ [a] ”は発信元MAC アドレス、第3番目のフィールド“ [B] ”は宛先IPアドレ

ス、第4番目のフィールド“[A]”は発信元IPアドレス、第5番目のフィールド“[データ]”は送信データである。

【0018】[S78] ここで通信アダプタ11に障害が発生したとする。

[S79] ホスト10は、通信アダプタ11の障害を検知し、バックアップ用通信アダプタ12への切り替え処理を開始する。

【0019】[S80] ホスト10は、バックアップ用通信アダプタ12の活性化を指示する。これにより、活性化された通信アダプタ12が、通信アダプタ11に代わってホストの通信機能を制御するようになる。

【0020】[S81] ホスト20に対してデータを送るため、ホスト10は、通信アダプタ12へその送信を依頼する。

[S82] バックアップ用通信アダプタ12は、ホスト10から供給されたデータにアドレス情報を次のように付加して送信する。

【0021】
“[b] [c] [B] [A] [データ]”

[S83] ホスト10から送られてきたメッセージに回答して、ホスト20は何等かの処理を行なう。ホスト20は、その処理結果をホスト10に送信する。この時点は、ホスト20および通信アダプタ21は、通信アダプタ1aが故障したことを知らないの、ホスト20のネットワークテーブル20aには依然として(b|a:A)が設定されたままである。したがってホスト10宛のメッセージの宛先MACアドレスは“a”となっている。

【0022】
“[a] [b] [A] [B] [データ]”

[S84] しかし通信アダプタ11は障害によって停止しているので、このメッセージを受信できない。結局ホスト20から発信されたこのメッセージはホスト10届かず、破棄される。

【0023】[S85] このような通信不能状態はホスト20のネットワークテーブル20aが正しい値に更新されない限り続く。しかし都合の悪いことに、ネットワークテーブル20aは、データ送信メッセージやARP要求などを実際に受信しない限り更新されない。

【0024】一方、TCP/IPプロトコルの規定により、このようなネットワーク上のアクティビティが20分間途絶えた場合、ネットワークテーブル中の関連するエントリは無効化される。結局のところ、通信不能状態が20分間続いた後、ネットワークテーブルの該当エントリを無効化する機構が働く。すなわちネットワークテーブルのタイムアウトである。こうなればはじめて、通信アダプタ11のMACアドレスのエントリがネットワークテーブルから除去されることになる。

【0025】[S86] 20分以上の時間が経過した後、ホスト10が再びホスト20にメッセージを送るものと仮

定する。通信アダプタ12は、データの送信依頼をホスト10から受ける。

【0026】[S87] この時点では、通信アダプタ12のネットワークテーブル12aのうち通信アダプタ21のMACアドレスに関するエントリはない。従って、通信アダプタ12は、メッセージの宛先のMACアドレスを取得すべく次のARP要求を同報モードで送信する。ホスト10から受けたデータは通信アダプタ12が保留する。

【0027】“[同報] [c] [B] [A]”

[S88] 宛先の通信アダプタ21がこの同報メッセージを受信する。通信アダプタ21は、ホスト20に受信したメッセージを渡すと共に、自分のネットワークテーブルにホスト10のアドレスを次のように登録する。すなわち、IPアドレス“A”とMACアドレス“c”との対応関係を表すエントリ(b|c:A)がネットワークテーブル20aに登録される。

【0028】[S89] ホスト20は通信アダプタ21を経由して、そのMACアドレス“b”を知らせるためにARP応答メッセージをホスト10に返す。

[S90] 通信アダプタ12がARP応答メッセージを受信し、自己のネットワークテーブル12aにエントリ(c|b:B)に登録する。

【0029】[S91] 通信アダプタ12は、ステップS87で保留しておいたデータに、MACアドレス、IPアドレスを下記のように付加したメッセージを送信する。

“[b] [a] [B] [A] [データ]”

以上のように、従来のシステムでは、ステップS85で20分間の時間が経過するまで、障害の発生した現用通信アダプタ11のMACアドレスが他の装置のネットワークテーブルに登録されていた。

【0030】

【発明が解決しようとする課題】従来のシステムでは、ステップS85で20分間の時間が経過するまで、障害の発生した現用通信アダプタ11のMACアドレスが他の装置のネットワークテーブルに登録されたままになっているため、この間ホスト20はバックアップ用通信アダプタ12に対してデータを送ることができない。ホスト10が障害を検知して、バックアップ用通信アダプタ12への切り替えを行なっても、他のホスト内のネットワークテーブルが更新されないの、通信を維持することができないという問題があった。TCP/IP通信において、コネクション型のプロトコルを使用するアプリケーションの中には、この20分の通信中断によって処理がタイムアウトするものがあった。

【0031】このように、ホスト10において信頼性を上げるために通信アダプタを二重化しても、従来のシステムでは効果的に働かないという問題があった。

【0032】

【課題を解決するための手段】本発明の目的は、この問題を解決するため、TCP/IPプロトコル等による通信を切

断することなく別の通信アダプタを使用して瞬時に切り替えを完了し、通信を継続することのできる通信アダプタの切替方法と切替方式を提供することである。

【0033】上記目的を達成するために、第1のホストの通信機能が第1と第2の通信アダプタによって二重化され、複数のホスト間の通信をTCP/IPプロトコルを使用して行なうネットワークシステムで、第1の通信アダプタから第2の通信アダプタに切り替える切替方法が提供される。

【0034】この切替方法によれば、第1のホストが第1の通信アダプタの故障を検出した時に、第1のホストは、第1の通信アダプタを切り離し、第2の通信アダプタを活性化する。切替えに際して、第1のホストはネットワークテーブル(ARPキャッシュ)を初期化し、第2のホストへの制御メッセージの送信を起動する。好ましくは、この制御メッセージはICMPプロトコルのエコー要求メッセージである。

【0035】第2の通信アダプタは、ARP 要求メッセージを同報送信し、第2のホストからARP 応答メッセージを受信する。第2の通信アダプタは、ARP 応答メッセージに含まれるアドレス情報に基づいて、前記制御メッセージを第2のホストに送信し、前記制御メッセージに対する第2のホストからの応答メッセージを受信する。

【0036】これによって第1のホストは、第1の通信アダプタから第2の通信アダプタへの切り替えが終了したと認識し、通信を再開する。このような切替方法は、第1のホストがトランスポート層の通信プロトコルを制御し、第1と第2の通信アダプタはインターネット層以下の通信プロトコルを制御するようなシステムに適用される。

【0037】また、上記目的を達成するために、第1のホストの通信機能が第1と第2の通信アダプタによって二重化され、複数のホスト間の通信をTCP/IPプロトコルを使用して行なうネットワークシステムで、第1の通信アダプタから第2の通信アダプタに切り替える切替方法が提供される。

【0038】この切替方法によれば、第1のホストが第1の通信アダプタの故障を検出した時に、第1の通信アダプタを切り離し、第2の通信アダプタを活性化する。続いて第1のホストから第2の通信アダプタを経由して第2のホストにARP 応答メッセージを送信する。このARP 応答メッセージは、その送信元アドレスフィールドにホストのIPアドレスと第2の通信アダプタのMAC アドレスを含んでいるので、ARP 応答メッセージを受け取った第2のホストはネットワークテーブルをこれに合わせて更新する。このような切替方法は、ホストがインターネット層の通信プロトコルを制御し、第1と第2の通信アダプタはネットワークインタフェース層の通信プロトコルを制御するシステムに適用される。

【0039】

【発明の実施の形態】図1から図6を用いて本発明の実施の形態を説明する。まず初めに、本発明の実施の形態における通信プロトコルの構造について述べる。図1は、TCP/IP プロトコルと、それに対応する通信アダプタ階層構造を説明する図である。

【0040】図1(A)は、TCP/IPプロトコルの機能階層をISO OSI(International Standards Organization - Open System Interconnect)参照モデルと比較して示す。下位層から順に、ネットワークインタフェース層(ISOのOSI 参照モデルでは物理層、データリンク層に相当)、インターネット層(OSI参照モデルではネットワーク層)、トランスポート層があり、この上に実際のアプリケーションがある。

【0041】ネットワークインタフェース層では、MAC(Media Access Control)アドレスによってパケットの宛先を指定する。インターネット層のプロトコルはIPプロトコルと呼ばれるもので、メッセージのヘッダにはIPアドレスが付加される。このIPアドレスに関連するプロトコルとして、IPアドレスからMAC アドレスへの変換を行なうARP(Address Resolution Protocol)、その逆の変換を行なうRARP(Reverse Address Resolution Protocol)がある。またネットワーク制御用プロトコルとして、ICMP(Internet Control Message Protocol)がある。トランスポート層プロトコルとしては、TCP(Transmission Control Protocol)とUDP(User Datagram Protocol)が提供される。これらのプロトコルはデータ通信分野においてよく知られているものであるので、詳細については省略する。

【0042】本発明で使用される通信アダプタは、プロトコル機能の実装方法によって二つのタイプに分類される。図1(B)のタイプでは、ホストはトランスポート層以上機能を担当し、通信アダプタはインターネット層以下の機能を担当する。言い換えると、通信アダプタはルータとしての機能を有する。したがって通信アダプタがIPアドレスとMAC アドレスの両方を管理する。

【0043】一方、図1(C)に示すタイプでは、通信アダプタは最下位層のみを担当し、プロトコルの大半はホスト側の通信ソフトウェアで制御する。この場合、IPアドレスはホストが管理する。

【0044】本発明の第1の実施の形態は、図1(B)のタイプの通信アダプタを二重化した場合の切り替え方法に関するものである。次のセクションでは図2と図3を参照して、第1の実施の形態を説明する。また本発明の第2の実施の形態は、図1(C)のタイプの通信アダプタを二重化した場合の切り替え方法に関するものである。これは図4と図5を参照して後述する。

【0045】図2は、本発明の第1の実施の形態の動作原理を説明するシーケンス図である。ここでは、図7の構成と同様に二つのホスト10、20がそれぞれの通信アダプタを通してネットワーク100でつながっている。上で

触れたように、第1の実施の形態では通信アダプタ11, 12がネットワークテーブル11a, 12aにてIPアドレスとMACアドレスの対応を管理する。ただし、ホスト20側は、先に説明した図8と同様にホスト20自身がネットワークテーブル20aを管理するものとする。実際のところ、第1の実施の形態の動作は、ホスト20のプロトコルスタックがどのように実装されているかには影響されない。

【0046】このような構成においてホスト10は次のようにして障害に対処する。

【S1】ホスト10は通信アダプタ11を使用してホスト20と通信を行なっている。

【0047】【S2】ここで通信アダプタ11に障害が発生する。

【S3】ホスト10は、この通信アダプタ11の障害を検知する。

【S4】ホスト10は、バックアップ用通信アダプタ12を通信アダプタ11の代替として割り当てる。

【0048】【S5】ホスト10は、IPアドレス登録要求を行う。

【S6】ホスト10は、ACT指示(活性化指示)を割り当てた通信アダプタ12に送る。

【0049】【S7】ACT指示に呼応して、通信アダプタ12はIPアドレスの登録とネットワークテーブルのクリアを行う。

【S8】通信アダプタ12は、ACT完了をホスト10に対して返答する。

【0050】【S9】ホスト10は、現在同一ネットワーク上で通信を行なっている相手を洗い出す。具体的には、ホスト10は内部の管理テーブルを参照して、現在自分がTCPコネクションを張っているノードを抽出する。

【0051】【S10】ホスト10は、洗い出された相手(ここではホスト20)に対して、制御メッセージを送信すべく、通信アダプタ12に送信依頼を行なう。この制御メッセージは、望ましくはICMPプロトコルに定義されているエコー要求メッセージである。このメッセージは宛先のホストがネットワークに正常に接続されているかどうかを確認するための制御メッセージであり、受け取ったホストはエコー応答メッセージによって、これに返答する。

【0052】【S11】ネットワークテーブルがクリアされているので、送信依頼に対応して、通信アダプタ12はARP要求を同報送信する。

【S12】ARP要求を受信したホスト20の通信アダプタ21が、発信元のMACアドレスが変更されたことを認識し、ネットワークテーブルに新しいアドレスを登録する。

【0053】【S13】ホスト20は通信アダプタ21を経由して、ARP応答メッセージを返す。

【S14】ARP応答を受信した通信アダプタ12は、得られたアドレス情報をネットワークテーブル12aに登録す

る。

【0054】【S15】通信アダプタ12は、エコー要求メッセージを送信する。

【S16】エコー要求メッセージに返答して、通信アダプタ21がエコー応答メッセージをホスト10に返す。

【0055】【S17】通信アダプタ12が受信したエコー応答メッセージをホスト10に渡す。

【S18】以上のステップによって、ホスト10、20間の通信に必要なアドレス情報が更新されたので、障害を起こした通信アダプタから切り替えられた通信アダプタ12を通じて両者間の通信が継続される。

【0056】以上の手順により、現用通信アダプタ11に障害が発生した時にバックアップ用通信アダプタ12に即時に切り替えてデータ通信を継続することが可能となる。従来、ネットワークテーブルを更新するのに必要であった20分間の待ち時間は解消される。

【0057】図3は、本発明の第1の実施の形態の詳細な動作を説明するシーケンス図である。これは、前述の図2に対応するものであって、MACアドレス、IPアドレスを以下のように具体的に設定して説明する。

【0058】A:ホスト10のIPアドレス

B:ホスト20のIPアドレス

a:現用通信アダプタ11のMACアドレス

c:バックアップ用通信アダプタ12のMACアドレス

b:ホスト20の通信アダプタ21のMACアドレス

図3において通信アダプタの切替プロセスは以下のステップに従って行なわれる。

【0059】【S21】ホスト10に接続された通信アダプタ11は正常に動作している。ホスト10は、通信アダプタ11に下記のホスト20宛のメッセージの送信を依頼する。

"[B][A][データ]"

ここで、先頭のフィールド"[B]"はホスト20のIPアドレスを表し、第2番目のフィールド"[A]"は発信元のIPアドレスを表し、第3番目のフィールド"[データ]"は送信データの内容を表す。

【0060】【S22】通信アダプタ11のネットワークテーブルには宛先IPアドレス"B"に対応するMACアドレスがまだ登録されていないので、通信アダプタ11は一旦このデータ送信を保留し、相手のMACアドレスを調べるために次のARP要求メッセージを送信する。

【0061】"[同報][a][B][A]"

ここで、先頭のフィールド"[同報]"は全てのMACアドレスに送信することを表すコードである。第2番目のフィールド"[a]"は発信元MACアドレス、すなわち現用通信アダプタ11のMACアドレスである。第3番目のフィールドは宛先IPアドレスを表し、第4番目のフィールドは発信元IPアドレスを表す。

【0062】【S23】バックアップ用通信アダプタ12は待機中ではあるが、ネットワークのアクティビティを

常時監視しており、アドレス更新に関する情報があれば、それを取り込むように構成されている。従って通信アダプタ12は、ステップS22で送信されたARP要求の情報を取り込み、自分自身のネットワークテーブルを構築する。ここではホスト10のIPアドレス“A”とMACアドレス“a”の対応関係がネットワークテーブル12aに登録される。このテーブルエントリを次のように表わす。

【0063】(c|a:A)

最初の“c”はMACアドレス“c”の装置のネットワークテーブルを意味する。2番目の“a”は受信したMACアドレスを、また3番目の“A”はこれに対応するIPアドレスを意味している。

【0064】[S24] 通信アダプタ21は、ステップS22で送信されたARP要求を受信し、自分自身のネットワークテーブル20aにホスト10のMACアドレス“[a]”とIPアドレス“A”の対応関係を表わすエントリ(b|a:A)に登録する。

【0065】[S25] ホスト20は自分宛のARP要求であることを認識して、通信アダプタを経由してARP応答メッセージをホスト10に返信する。

[S26] 通信アダプタ11はARP応答を受信し、ネットワークテーブルにホスト20のMACアドレス“[b]”とIPアドレス“B”の対応関係を表わすエントリ(a|b:B)に登録する。

【0066】[S27] 宛先MACアドレスを得ることができたので、通信アダプタ11はステップS22で保留しておいたデータをホスト20に向けて送信する。この時のメッセージには下記のようにIPアドレス、MACアドレスを設定する。

【0067】

“[b] [a] [B] [A] [データ]”

ここで、先頭のフィールド“[b]”は宛先MACアドレス、第2番目のフィールド“[a]”は発信元MACアドレス、第3番目のフィールド“[B]”は宛先IPアドレス、第4番目のフィールド“[A]”は発信元IPアドレス、第5番目のフィールド“[データ]”は送信データである。

【0068】[S28] ここで現用通信アダプタ11に障害が発生したとする。

[S29] ホスト10は、通信アダプタ11の障害を検知し、バックアップ用通信アダプタ12への切り替え処理を開始する。

【0069】[S30] ホスト10は、バックアップ用通信アダプタ12の活性化を指示する。これに呼応して通信アダプタ12は自分のネットワークテーブルをクリアし、AC T完了をホスト10に対して返答する。これにより障害を起こした通信アダプタ11からバックアップ用通信アダプタ12への切り替えが完了する。

【0070】[S31] ホスト10は、ホスト20に対して、次のようなメッセージを送信すべく、通信アダプタ

12に送信依頼を行なう。

“[B] [A] [エコー要求]”

このメッセージはICMPプロトコルに定義されているエコー要求メッセージであり、本来は宛先の装置が動作可能状態であるかどうかを問い合わせるためのメッセージである。

【0071】[S32] 通信アダプタ12のネットワークテーブルはステップS30で初期化されたため、通信先のIPアドレスに対応するMACアドレスが登録されていない。そこで、通信アダプタ12は下記のARP要求を通信相手に送信する。

【0072】“[同報] [c] [B] [A]”

[S33] ステップS32のARP要求を受信した通信相手のホスト20が自己のネットワークテーブル中のIPアドレス“A”に関するエントリを更新する。その結果、アドレスの対応関係を表わすエントリ(b|c:A)が登録される。

【0073】[S34] 通信アダプタ21は通信アダプタ12に対してARP応答を返し、要求されたMACアドレスが“b”であることを示す。

“[c] [b] [A] [B]”

[S35] ARP応答を受信した通信アダプタ12はネットワークテーブルにホスト20に関するアドレスの対応関係を表わすエントリ(c|b:B)に登録する。

【0074】[S36] 通信アダプタ12は、ステップS32で保留したICMPエコー要求メッセージを、ネットワークテーブル12aに登録されたアドレス情報を使用してホスト20宛に送信する。

【0075】

“[b] [c] [B] [A] [エコー要求]”

ここで、先頭のフィールド“[b]”は通信相手の通信アダプタ21のMACアドレスであり、第2番目の“c”は発信元のバックアップ用通信アダプタ12のMACアドレスである。第3番目のフィールド“[B]”は通信相手のホスト20のIPアドレスであり、第4番目のフィールド“[A]”はホスト10のIPアドレスであり、第5番目の“[エコー要求]”は送信依頼を受けたデータである。

【0076】このように発信元のMACアドレスがステップS27の場合のMACアドレス“a”からMACアドレス“c”に変わっている。これはホスト10側の通信アダプタが正しくバックアップ用通信アダプタ12に切り変わったことを示している。

【0077】以上のように、現用通信アダプタ11に障害が発生した場合に、バックアップ用通信アダプタ12に即時に切り替え、データ通信を継続することが可能となった。また以上の説明からわかるように、このような通信アダプタの切替制御および切り替えに伴うアドレス情報の更新制御はホスト10が主導して行なうものである。したがってホスト10は、これらの制御のためのプログラムを備えている。またこの制御プログラムは、コンピュー

タが読み取ることのできる媒体によって配布されるものである。

【0078】次に、本発明の第2の実施の形態について述べる。第2の実施の形態はホスト10の通信アダプタが下位層の通信機能のみを担当する構成となっている点において、第1の実施の形態と異なっている(図1(C)を参照)。この構成では、ホスト10がIPアドレスとMACアドレスの対応を管理するネットワークテーブル10aを有する。

【0079】図4は、本発明の第2の実施の形態の動作原理を説明するシーケンス図である。両ホストは以下のステップに従って通信を行なう。

【S41】ホスト10は通信アダプタ11を使用してホスト20と通信を行なっている。

【0080】【S42】通信アダプタ11に障害が発生する。

【S43】ホスト10は、この通信アダプタ11の障害を検知する。

【S44】ホスト10は、待機装置の割り当てを行う。ここでは、バックアップ用通信アダプタ12を通信アダプタ11の代替機として割り当てる。

【0081】【S45】ホスト10は、ACT 指示(活性化指示)を割り当てた通信アダプタ12に送る。

【S46】通信アダプタ12がACT 完了をホスト10に返答する。これにより、バックアップ用通信アダプタ12が動作し得る状態になる。

【0082】【S47】ACT 完了を受けると、ホスト10はネットワークテーブル10aを更新して、MAC アドレスをa からc に変更する。

【S48】ホスト10は、現在同一ネットワーク上で通信を行なっている相手を洗い出す。具体的には、ホスト10は内部の管理テーブルを参照して、現在自分がTCP コネクションを張っているノードを抽出する。

【0083】【S49】ホスト10は、ネットワーク上の他のノードのネットワークテーブルの更新を促すためにARP 応答メッセージを送信する。例えば、ホスト20の通信アダプタ21は、このARP 応答メッセージを受信すると、ネットワークテーブルのホスト10に関するエントリを更新し、ホスト10のIPアドレスと通信アダプタ12のMAC アドレスの新しい対応関係(b | c:A)を登録する。

【0084】【S50】以上のステップによって、ホスト10、20間の通信に必要なアドレス情報が更新されたので、障害を起こした通信アダプタから切り替えられた通信アダプタ12を通じて両者間の通信が継続される。

【0085】以上の手順により、現用通信アダプタ11に障害が発生した時にバックアップ用通信アダプタ12に即時に切り替えてデータ通信を継続することが可能となる。従来、ネットワークテーブルを更新するのに必要であった20分間の待ち時間は解消される。第2の実施の形態ではTCP/IPの上位層をホスト10が制御しているの

で、通信アダプタの切り替え直後にARP 応答メッセージを直接発信することができる。

【0086】図5は、本発明の第2の実施の形態の詳細な動作を説明するシーケンス図である。これは、前述の図4に対応するものであって、MAC アドレス、IPアドレスを第1の実施の形態と同様に設定して説明する。

【0087】【S51】ホスト10に接続された通信アダプタ11が正常に動作している状況を仮定する。ホスト10は、下記のホスト20宛のメッセージの送信しようとしている。

【0088】" [B] [A] [データ] "

【S52】ホスト10のネットワークテーブル10a には宛先IPアドレス"B" に対するMAC アドレスがまだ登録されていないので、ホスト10は一旦このデータ送信を保留し、相手のMAC アドレスを調べるために次のARP 要求メッセージを同報送信する。

【0089】" [同報] [a] [B] [A] "

【S53】ARP 要求メッセージを通信アダプタ21が受信してホスト20に渡す。ホスト20は、IPアドレスとMAC アドレスの対応関係を表わすエントリ(b | a:A)をネットワークテーブル20a に登録する。

【0090】【S54】ホスト20は自分宛のARP 要求であることを認識して、ホスト10に対して下記のARP 応答を返信する。

" [a] [b.] [A] [B] "

【S55】通信アダプタ11はARP 応答を受信しホスト10に渡す。ホスト10は、ホスト20のMAC アドレスとIPアドレスの対応関係を表わすエントリ(a | b:B)をネットワークテーブル10a に登録する。

【0091】【S56】宛先MAC アドレスを得ることができたので、ホスト10はステップS52で保留しておいたデータをホスト20に向けて送信する。

" [b] [a] [B] [A] [データ] "

【S57】しかし、ここで現用通信アダプタ11に障害が発生したとする。

【0092】【S58】ホスト10は、通信アダプタ11の障害を検知する。

【S59】ホスト10は、バックアップ用通信アダプタ12の活性化を指示する。

【S60】通信アダプタ12は活性化完了メッセージを返して、準備ができたことをホスト10に知らせる。

【0093】【S61】ホスト10は、自分自身のネットワークテーブル10a を更新し、バックアップ用通信アダプタ12のものにする。すなわち(a | b:B)から(c | b:B)に更新する。

【0094】【S62】ホスト10は、MAC アドレスが"a" から"c" に更新されたことをホスト20に知らせるために、下記のARP 応答メッセージを送信する。

" [b] [c] [B] [A] "

【S63】ARP 応答メッセージを受信した通信アダプタ

21はそれをホスト20に渡す。ホスト20はネットワークテーブル20aを更新し、旧エントリ(b|a:A)を(b|c:A)に書き替える。

【0095】[S64]ホスト10がホスト20宛の下記のメッセージの送信を通信アダプタ12に依頼する。

“ [B] [A] [データ] ”

[S65] この依頼に回答して、通信アダプタ12は下記のメッセージをホストに送信する。

【0096】

“ [b] [c] [B] [A] [データ] ”

以上のように、現用通信アダプタ11に障害が発生した時、ホスト10はバックアップ用通信アダプタ12に切り替えると共に、ARP 応答を同報送信してネットワークテーブルを更新させる。これによってデータの送受信を中断せずに続けることが可能となる。

【0097】また第1の実施の形態と同様に、ホスト10は、このような通信アダプタの切替制御および切り替えに伴うアドレス情報の更新制御のためのプログラムを備えている。この制御プログラムは、コンピュータが読み取ることのできる媒体によって配布される。

【0098】次に、図6を参照して、本発明の第3の実施の形態の構成を示す。図6のシステムは、グローバルサーバ40に同一の通信のネットワーク100に複数の通信アダプタ41ないし45を接続して負荷を分散させると同時に、複数のバックアップ用装置を使用したホットスタンバイシステムを実現したものである。例えば図6が示すように、通信アダプタ41ないし45にそれぞれMAC アドレスMACa、MACb、MACc、MACd、MACeを付与し、現用通信アダプタの41、42、43にはIPアドレスIPA、IPB、IPCをそれぞれ設定する。他の通信アダプタの44、45はバックアップ用である。

【0099】現用通信アダプタの41、42、43のいずれかに障害が発生した場合、バックアップ用通信アダプタ44、45のいずれかを代替として割り当てる。図1(B)のように通信アダプタがネットワーク層以下をサポートしている場合は、新しい通信アダプタにIPアドレスを設定すると共に、ネットワークテーブルをクリアして、ARP 要求を他のノードに送信する。これによって他のホストのネットワークテーブルが更新されるので、切り替えた通信アダプタを使用して通信を即時に再開することが可能となる。

【0100】また図1(C)のようにホストがネットワーク層以上をサポートしている場合は、ホストが管理するネットワークテーブルをクリアして、ARP 応答を他の

ノードに送信する。これによって他のホストのネットワークテーブルが更新されるので、切り替えた通信アダプタを使用して通信を即時に再開することが可能となる。このような図6の構成は、高い通信性能を要求されるとともに、高信頼性が必要なサーバ、いわゆるグローバルサーバに適している。

【0101】

【発明の効果】以上説明したように、本発明によれば、通信中の通信アダプタの障害を検知したときにその通信アダプタを切り離すと共に、ネットワークテーブルをリセットあるいは更新し、ARP 要求あるいはARP 応答を送信する。そして、これらのメッセージに含まれるアドレス情報をネットワークテーブルに登録して通信を継続する。このため、TCP/IPプロトコルによる通信を切断することなく予備の通信アダプタを使用して瞬時に切り替えを完了し、通信を継続することができる。

【図面の簡単な説明】

【図1】TCP/IPプロトコルと、それに対応する通信アダプタ階層構造を説明する図であり、(A)はTCP/IPプロトコルの階層構造をISO OSI 参照モデルと比較して説明する図であり、(B)は本発明の第1の実施の形態のホストと通信アダプタにおけるプロトコル制御の階層構造を説明する図であり、(C)は本発明の第2の実施の形態のホストと通信アダプタにおけるプロトコル制御の階層構造を説明する図である。

【図2】本発明の第1の実施の形態の動作原理を説明するシーケンス図である。

【図3】本発明の第1の実施の形態の詳細な動作を説明するシーケンス図である。

【図4】本発明の第2の実施の形態の動作原理を説明するシーケンス図である。

【図5】本発明の第2の実施の形態の詳細な動作を説明するシーケンス図である。

【図6】本発明の第3の実施の形態の構成を示す図である。

【図7】従来のホスト間通信システムの構成を示す図である。

【図8】従来技術の通信アダプタの切り替えを説明するシーケンス図である。

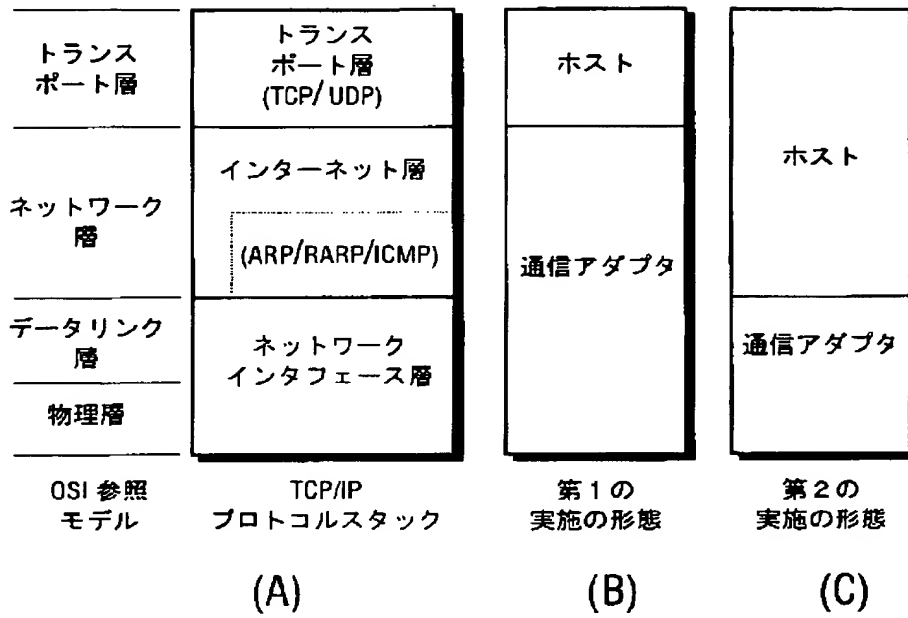
【符号の説明】

10、20 ホスト

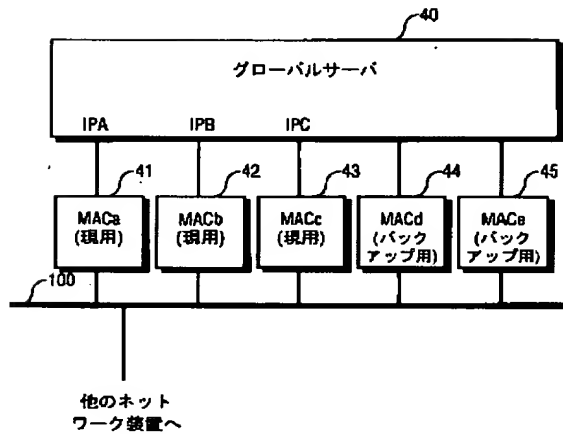
11、12、21 通信アダプタ

100 ネットワーク

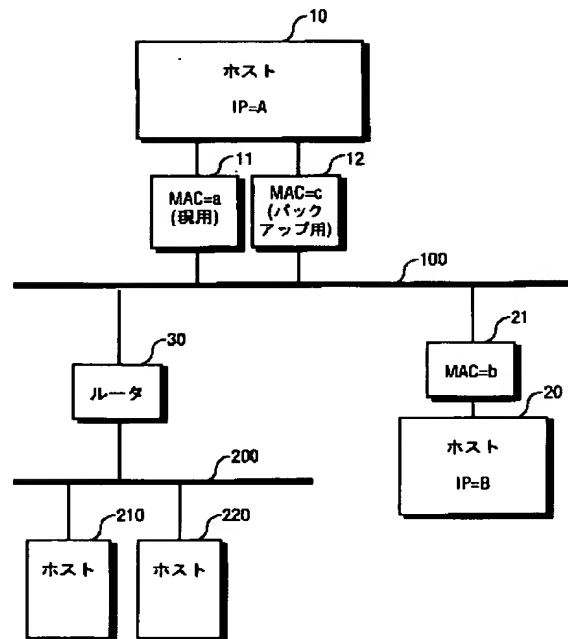
【図1】



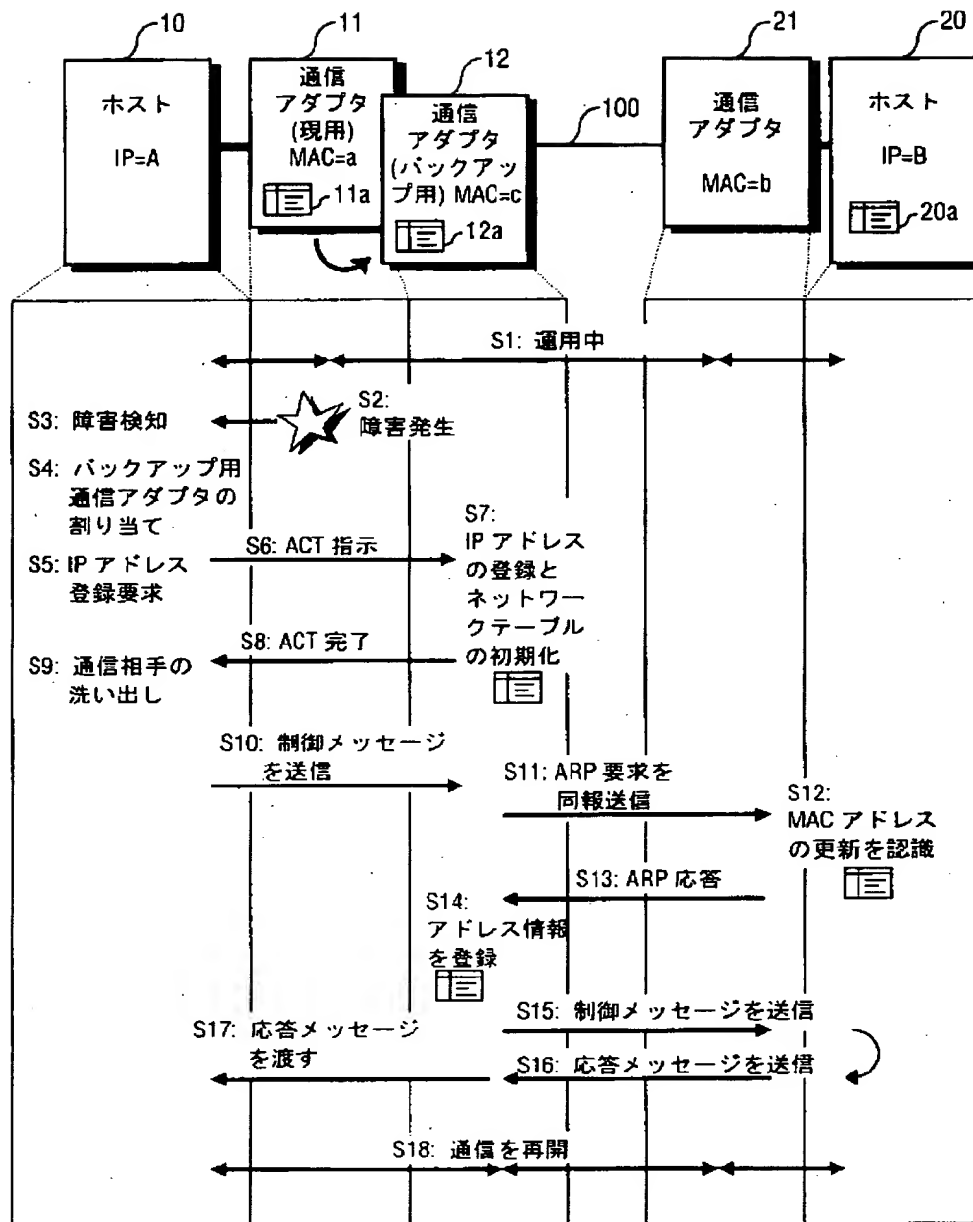
【図6】



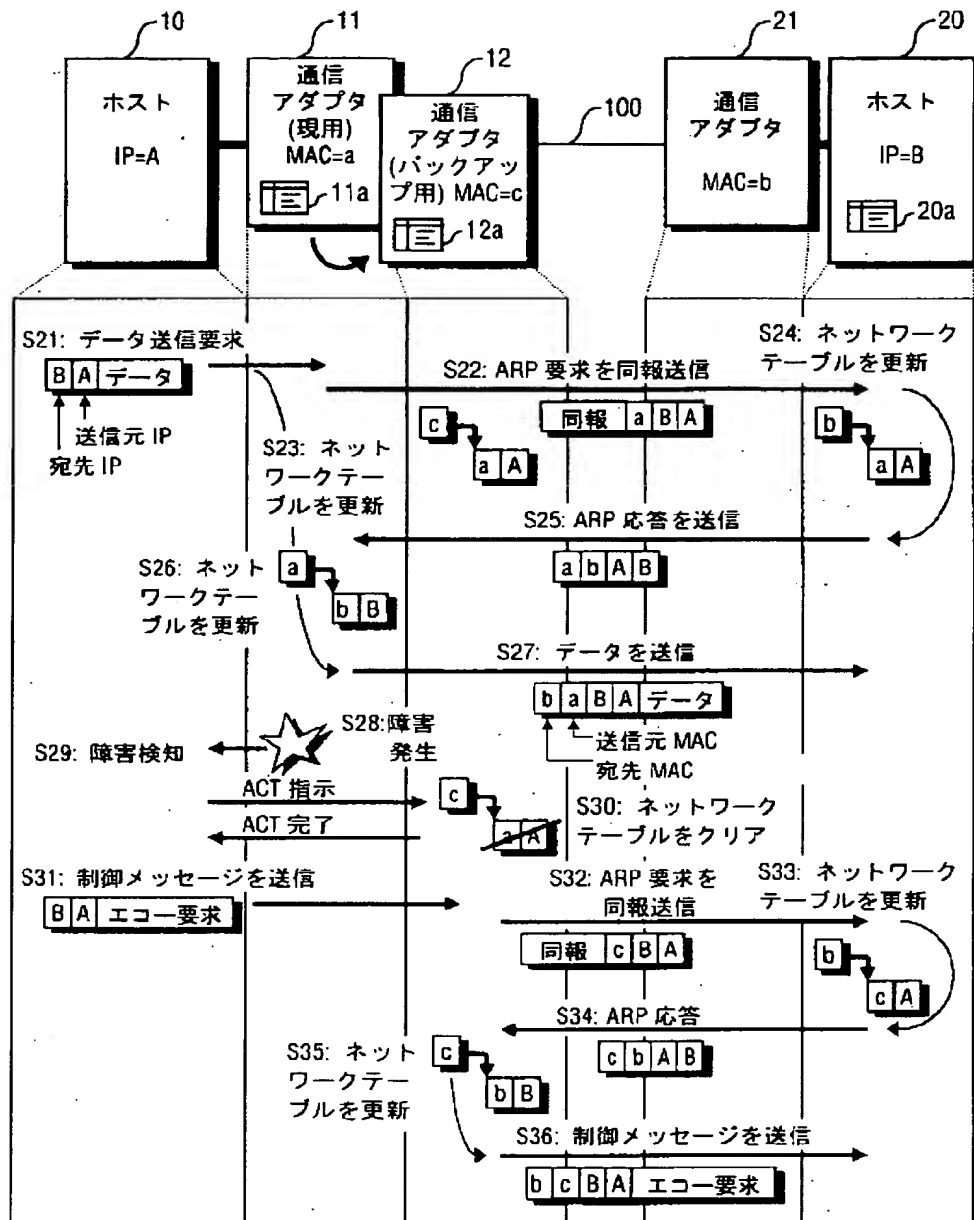
【図7】



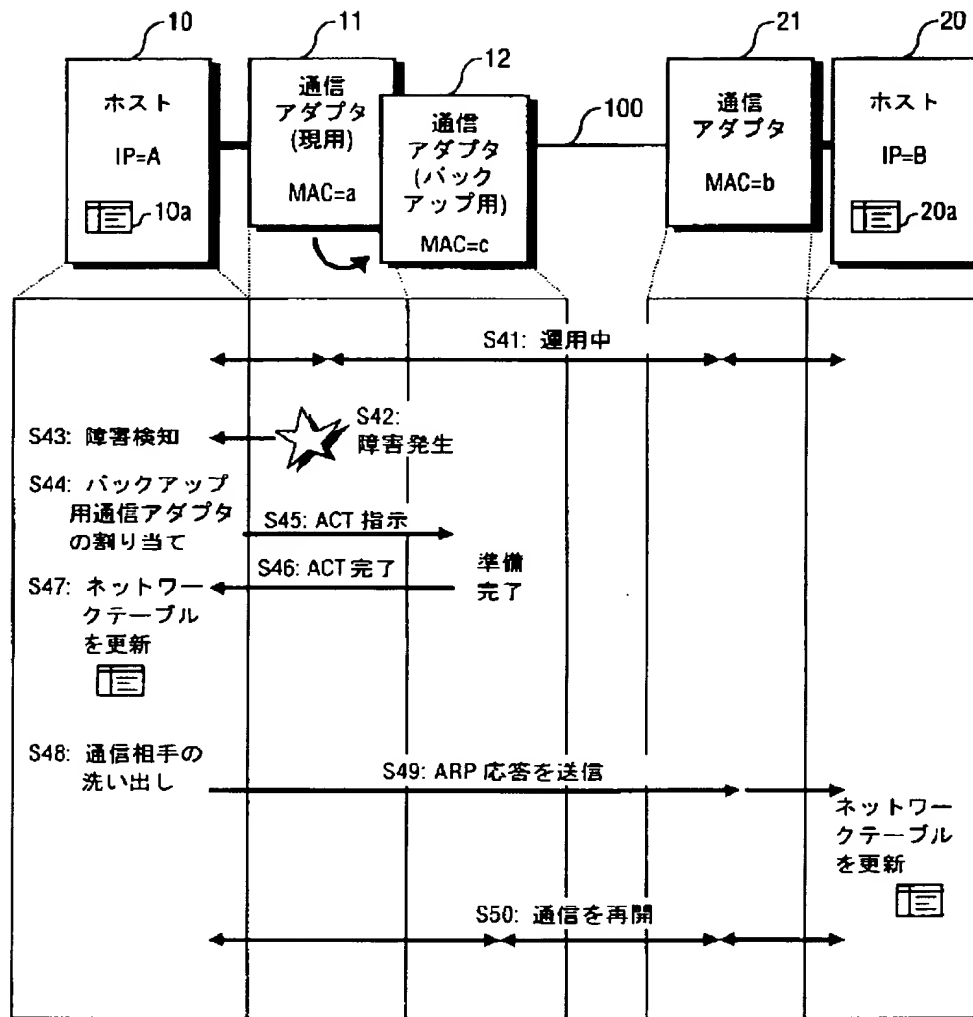
【図2】



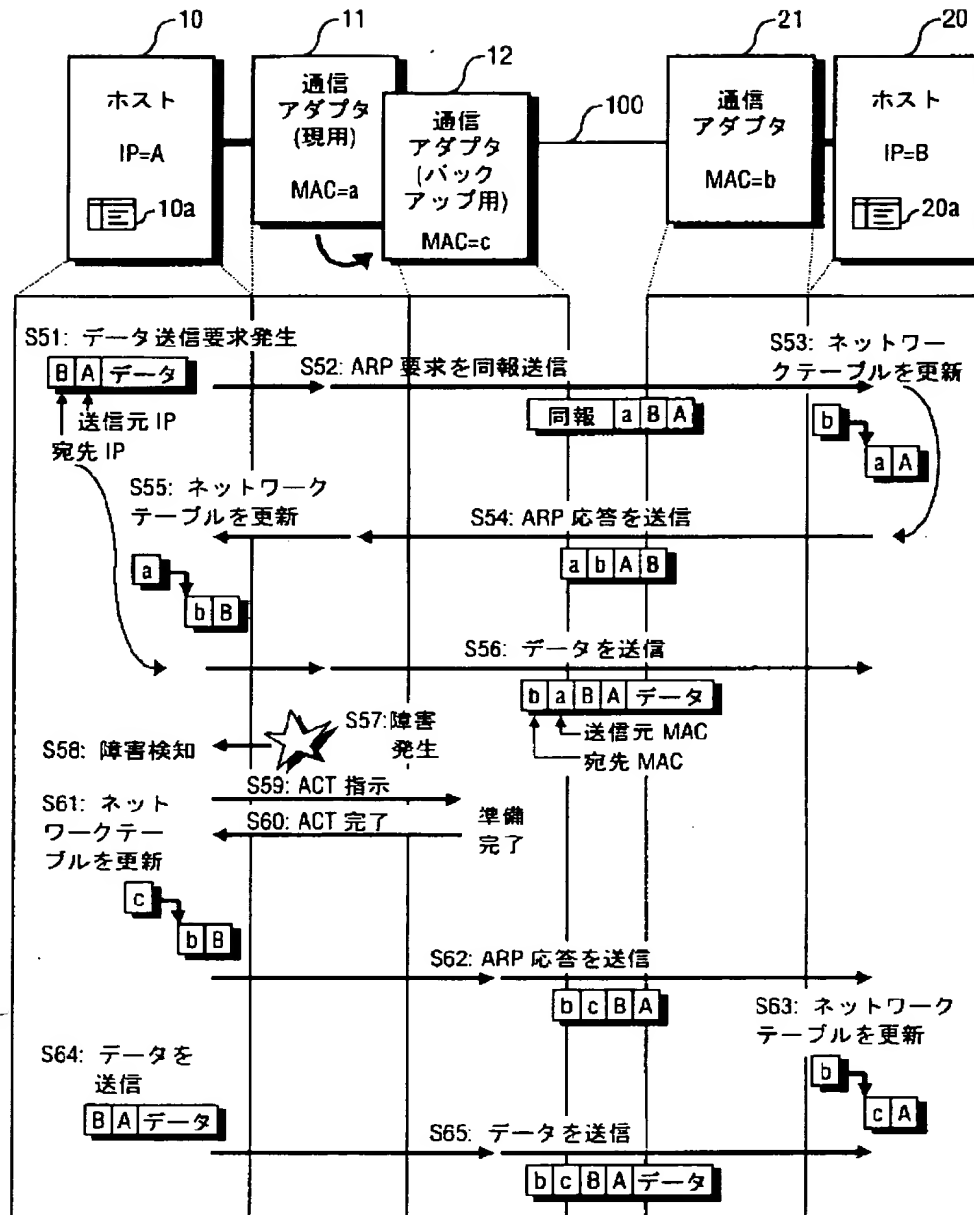
【図 3】



【図4】



【図5】



【図8】

